



**Supplementary figure S1.** Functional analysis of core and accessory genes. **(A)** distribution of GO functional terms among all the strains accessory genome as a variability measure. **(B)** Box-and-whisker plots of functional terms overrepresented in core genes compared to accessory genes among genomic clusters. The box plots were drawn with the upper and lower lines corresponding to the first and third quartiles, the middle lines corresponding to the medians, and the upper and lower whiskers extended to 1.5 times the interquartile ranges. Color dots represent each cluster value. Note the difference in scale between the panels.